Integrating gene expression with protein interactions for drug target prioritization

Griet Laenen 1,2, Lieven Thorrez 1,3,4, Daniela Börnigen 5, Yves Moreau 1,2

1) KU Leuven, Department of Electrical Engineering - ESAT, SCD-SISTA, Leuven, Belgium
2) iMinds Future Health Department, Leuven, Belgium
3) KU Leuven, Department of Development and Regeneration, Lab of Translational Cardiomyology, Leuven, Belgium
4) KU Leuven Kulak, Subfaculty of Medicine, Kortrijk, Belgium
5) Harvard University, Harvard School of Public Health, Biostatistics Department, Boston, MA, USA

E-mail: griet.laenen@esat.kuleuven.be

Many drugs achieve their effects by modulating multiple targets. These targets are, however, often unknown and difficult to find among the thousands of gene products. Yet this knowledge could be of substantial value to future drug development, in particular for side effect prediction and drug repositioning.

We propose a computational method tackling this target identification problem by the analysis of gene expression following drug treatment in the context of a functional protein association network. More specifically, genes are prioritized as potential targets based on the transcriptional response of functionally related genes. To this end, differential expression signals are diffused over the network either using a kernel-based random walk or on the basis of connectivity correlations between nodes.

Both diffusion strategies were evaluated on 235 publicly available gene expression datasets for treatment with bioactive molecules having a known target. With AUC values of 91% and 92% respectively, the best results were obtained by a single-step symmetric normalized Laplacian kernel diffusion of moderated t-statistics and by a correlation-based diffusion of log2 ratios.

The obtained AUC values indicate the predictive power of integrating experimental gene expression data with prior knowledge on protein interactions to identify the targets of a drug.